AN ULTRA-LOW POWER AND DISTRIBUTED ACCESS PROTOCOL FOR BROADBAND WIRELESS SENSOR NETWORKS

Lizhi Charlie Zhong, Rahul Shah, Chunlong Guo, Jan Rabaey
Berkeley Wireless Research Center
Department of EECS
University of California at Berkeley

Abstract - **An ultra-low power access control scheme has been proposed for broadband wireless sensor networks. This is a distributed method, which does not require the existence of a central base station, eliminating the possibility of a single point failure. This access protocol combines the best of simple carrier sense multiple access (CSMA) and spread spectrum techniques. It trades the bandwidth in broadband applications for higher power efficiency and throughput. Compared with other access schemes, it consumes far less power. This access protocol does not require a dedicated control channel, or synchronization, whether global or local. Additionally, it has very low delay and does not have the problem of coordinating broadcast and scheduled unicasts. For sensor networks, the inherent redundancy can be exploited to further improve power efficiency. It achieves much higher power efficiency than much more complicated algorithms seeking to reduce the redundancy.**

I. SENSOR NETWORKS

Sensor network has attracted quite a n attention recently in both academia and industry. The idea is simple enough. Tens or hundreds of sensors are spread out. Each of them can measure the temperature, light intensity, humidity, noise level, airflow speed, etc. in its locality. But the beauty of the sensor network is that all these nodes can be self -organized into a network, over which they can do cooperative processing to accomplish tasks that they cannot do individually.

The idea of sensor network applies directly to office building environm ent, where sensors can work together to create a microclimate based on personal settings. Similar things can be done inside a car, airplane, train, bus, or in any other public transportation and in any public places. But the applications for sensor network s go far beyond these. They can be used in smart home, in the battlefield, in scientific research, in home nurse care and interactive museums. Sensor networks can also be used together with Internet to create a net surfing experience impossible before. Loc al sensor networks can recreate the environment where a web page tries to bring one to, be it smells, strong wind or cold weather. Similarly, a game designer can put in all the details in a compute game, which allows sensors to recreate the scene in a play er's local environment. Most importantly, the "sensor" in a sensor network does not have to be limited to a mechanic device that measures certain parameters. It can be anything that has certain knowledge of its local environment, for example, an individual person can be a "sensor". Therefore, almost everything can be thought of as a form of sensor networks, so design used in a traditional sensor network can be ported to it easily.

Despite the diversity of the applications in sensor networks, the following requirements are true for all of them: low power, low cost, wireless and ad -hoc. Among these requirements, low power is the most important. It would be a maintenance nightmare having to replace the batteries of the sensors every day or even every month. In some applications, the sensors are buried inside a wall so it is impossible to take them out to replace their batteries. In yet some other applications, the sensors are used just once and live only as long as their batteries can support. So it is very imp ortant to make the sensors self powered or have battery life of several years. In either case, the design of the entire system has to be optimized for low power consumption.

II. MEDIA ACCESS CONTROL PROTOCOLS

There are many different Media Access Contro l (MAC) protocols in the literature. They can be classified into different categories based on different principles. Some are centralized, with the base station or group leader doing the access control; some are distributed. Some use a single channel; some use multiple channels. Some use various versions of random access; some use reservation and scheduling. They are also optimized for different things: power, delay, throughput, fairness, Quality of Service (Qos) or support for multiple services.

The MAC for sensor networks has to be distributed and optimized for power. A distributed method using multiple channels and random access is the way to go. Distributed methods do not require a central control point, avoiding the problem of a single point of failure . Global synchronization is typically not required in distributed methods. Using multiple channels reduces collisions and retransmissions. It also reduces delay and increases throughput. Random access does not require a node to have any knowledge of the ne twork, therefore synchronization is not needed and there is no overhead in exchanging schedules and reservation information. Protocols using reservation and scheduling have better channel utilization, but they only make sense for periodic events, therefore they are not suited for mobile or any dynamic environment.

But there are very few papers on low power distributed MAC design. Code Division Multiple Access (CDMA) is combined with Time Division Multiple Access (TDMA) in [1] and with random access in [2]. But both are centralized methods and are not optimized for power. Several useful power saving techniques were presented in [3][4] using centralized methods. Distributed methods using CDMA are described in [5][6], but they do not support sleep mode. That is, a node is not powered off when it is not receiving or transmitting. Another distributed method uses multiple channels to implement collision avoidance (CA) by assigning each sender or receiver a distinct code [7]. It does not support sleep mode either.

A distributed method that supports sleep mode was proposed in [8] for sensor networks. It uses distributed scheduling and CDMA. Only local synchronization is needed. The scheduling is done pairwisely. Each node has a super-frame structure, which contains a boot-up phase and a reservation phase. But the alignment of super-frames is not synchronized to a single time reference. Instead it starts when the very first appointment is made for the node, so the alignment might be different from node to node. In boot-up phase, a node is powered on to search for a partner to make an appointment. Through a 4-way handshake, an appointment is made between two nodes, which manage to talk together. Only relative time is used in scheduling since the clock of each node is independent. The time reference common to the two nodes is established when the first message in the 4-way handoff is received, ignoring propagation delay. Once the appointment is made, it is expected to be periodic even though the appointment can be updated or removed at the end of current appointment. Only one appointment is set up between any two nodes. Since each node goes by its own local schedule and relative time is used, global synchronization is not required. Multiple code channels are used to alleviate the potential problem that time slots of different pairs of nodes might overlap.

The disadvantage of this scheme is that it has long delay in a multi hop path, since each node in the path has to wait for its scheduled slot in every hop. It is very difficult to coordinate broadcast and scheduled unicasts. Broadcast cannot be scheduled; otherwise global synchronization will be required. This scheme also has significant overhead in power since a node has to stay powered on in every boot-up phase. The low cost oscillators used in sensors tend to have relatively large frequency variations. Using relative time will not fix this problem. Guard time is required or the nodes have to resynchronize themselves after several super-frames, both incurring power penalty. This scheme is only useful when the nodes need to talk to one another periodically; otherwise the overhead of scheduling is not justified. It is also possible that the boot-up windows of two neighboring nodes never overlap, so they cannot talk to each other directly, and have to go through a longer path to communicate, which means more power consumption. This protocol also does not exploit the inherent redundancy in the sensor networks.

III. PICO-RADIO PROJECT

Pico-radio network by Berkeley Wireless Research Center (BWRC) is designed to be an ultra-low power, wireless ad-hoc network [9][10]. The projected power consumption for each node is $100\mu W$ and cost is less than 50 cents. The size of each node is 1 cm^3. The power consumption is minimized vertically across different layers and horizontally through the entire network. That is, the power is not only optimized for a particular layer, but also for all layers. The power is not only optimized for a particular node, but also for the entire network.

The MAC protocol we propose combines the best of spread spectrum techniques and Carrier Sense Multiple Access (CSMA). It extends CSMA into multi-channel CSMA. A channel can be code, time slot, sub-carrier/frequency or space. Conventional wireless Ethernet [11] can be thought of as single-channel CSMA. Even though multiple channels are offered in its physical layer, its MAC layer still thinks there is only one channel. Our approach fully exploits the diversity available. Before transmission, a node will randomly select a channel to see if it is busy. If the channel is busy, the node will randomly select another channel from the remaining channels. This process will go on until the node finds a channel that is idle and uses it to transmit its message. If all channels are busy, the node will set a random timer for each of them and back off. It will use the channel whose timer expires first and clear the timers for all other channels. There is no central control point in this algorithm, so it is fully distributed. A node does not need to know about other nodes in the network and can operate independently. Synchronization, local or global, is not required.

Low power consumption is being achieved by saving power in every aspect. First, a simple algorithm like random access means less power in processing and no power wasted in handshaking. With sleep mode, the main radio is turned off whenever it is not doing anything. This results in dramatic power saving. The use of multiple channels reduces collision, therefore reduces power for retransmission.

To support sleep mode, a wakeup radio is used in the physical layer, which consumes less than $1\mu W$ running at full duty cycle. It wakes up the main radio and informs it of the channel it should tune to. Since the destination ID is modulated into the wakeup signal, only the destination node will be waken up. With this design, not only the main radio can run at very low duty cycle, i.e. be turned off whenever it is not receiving or transmitting anything, there is also no delay for a node in waiting for a slot. Due to the multiple channels available, there are few potential users for the same channel at the same time, so the delay is small and throughput is high. Our approach is very simple and therefore very robust. It does not have all the overhead of the scheduling algorithm, nor the potential problems associated with it.

For sensor networks, the inherent redundancy can be exploited to further improve power efficiency. It is shown later that it achieves much higher power efficiency than very complicated algorithms that seek to reduce redundancy directly.

IV. EXPLOITING THE REDUNDANCY

In sensor networks, the same type of data from neighboring sensors is usually highly correlated. So there is quite a redundancy in the network. There are two ways to take advantage of this property. One way is to reduce or even eliminate the redundancy. For example, if both sensor A and B want to send data to the same destination, and their data are highly correlated, each sensor can transmit fewer bits; the receiver will make a joint decision based on both data and the correlation between them. This is called distributed source coding [12]. There is currently no constructive way to do this. Even if one is found in the future, very complicated signal processing will be needed to jointly decode the data. The benefit of all these complexities is the saving of only one or two bits per data packet. Another way is to allow the redundancy built in the network, and the destination will recover the lost data packets based on other data it obtains from the neighborhood of the original sender. Due to the high correlation between the data of a sensor and the data from its neighboring sensors (spatial correlation), and the high correlation between the current data and past data of the same sensor (temporal correlation), it is possible for the receiver to recover the original information from the vast data available which are highly correlated to the original data.

With this approach, acknowledgement can be eliminated. Simple error detection code can be used at the transport level to allow the destination to detect the error, or destination can periodically check the data, and if it misses a data packet in a cycle, it will trigger the estimation process to recover the lost data packet. It is shown in the next section the data estimation is very simple. By removing the acknowledgement at link level, we can save at least 304 bits for every data packet!

To obtain the data highly correlated to the original data, the first place to look is the destination itself. In this case, no transmission power will be wasted in data exchange. In some applications, where all the sensor data are sent to a monitor, it is reasonable to assume the monitor has all the data from all the sensors in a local neighborhood. So the monitor only needs to filter out the highly correlated data based on the location of the data source and the time when the data were sampled. A window size can be chosen for both geographic range and the time span of the usable data. So only data highly correlated to the original data will be used to do the data mining, which also reduces the complexity of the estimation filter. In networks where all nodes have similar set of functionalities, the destination doesn't necessarily have all the data to recover the original data. So it needs to get them from the network. It will first try to get them from its local neighbors. Only when there is not enough highly correlated data in its local neighborhood, which is very unlikely, it will slowly increase the network span, for example, first one hop away, then 5 hops away, then 10 hops away …

V. DATA ESTIMATION

Without loss of generality, the case for two sensors is considered. The same method can be extended to arbitrary number of sensors with minor modifications. Let's define the data from the source to be X, and the data from its neighbor Y. Both X and Y are random processes, and their samples are highly correlated. In case of an error, the destination does not receive X_n, but X_{n-i} and Y_{n-j} are available, where $i, j = 1, \cdots, p$. All the available data samples can be denoted by vector W_n. So the destination needs to estimate X_n given W_n. This is a classical estimation problem. The minimum mean square error estimate of X_n given W_n turns out to be $E[X_n \mid W_n]$. Linear estimate is used here to reduce the filter complexity at the cost of some performance loss. The problem now becomes finding the coefficients $\{a_i\}$ and $\{b_j\}$ such that the mean square error between $\hat{X}_n = \sum_{i=1}^{p} a_i X_{n-i} + \sum_{j=1}^{p} b_j Y_{n-j}$ and X_n is minimized.

The mean square error (MSE) is given by

$$\begin{aligned}\varepsilon_p &= E[e_n^2] = E[(\hat{x}_n - \sum_{i=1}^{p} a_i x_{n-i} - \sum_{i=1}^{p} b_i y_{n-i})^2] \\ &= \phi_{xx}(0) + \sum_{i=1}^{p}\sum_{j=1}^{p} a_i a_j \phi_{xx}(i-j) + \sum_{i=1}^{p}\sum_{j=1}^{p} b_i b_j \phi_{yy}(i-j) - 2\sum_{i=1}^{p} a_i \phi_{xx}(i) \\ &\quad - 2\sum_{i=1}^{p} b_i \phi_{xy}(i) + 2\sum_{i=1}^{p}\sum_{j=1}^{p} a_i b_j \phi_{xy}(i-j) \qquad (1)\end{aligned}$$

Differentiating w.r.t. a_i and b_j, and equating to zero, a set of equations analogous to the famous Yule-Walker equations is obtained.

$$\sum_{i=1}^{p} a_i \phi_{xx}(i-j) + \sum_{i=1}^{p} b_i \phi_{xy}(i-j) = \phi_{xx}(j) \qquad j = 1, \ldots, p \qquad (2)$$

$$\sum_{i=1}^{p} b_i \phi_{yy}(i-j) + \sum_{i=1}^{p} a_i \phi_{xy}(i-j) = \phi_{xy}(j) \qquad j = 1, \ldots, p \qquad (3)$$

Since both X and Y can be assumed to be wide sense stationary, auto-correlation and cross-correlation functions can be approximated by the corresponding time averages. The above derivations give the best linear estimator. But there are several special cases of this estimator, which are simpler and more useful. Only the previous sample of the source can be used to estimate the current sample. To make it even simpler, the value of the previous sample can be directly used as the estimate, eliminating the need for computing auto-correlation and buffering. Moving average can also be thought of as a special case of the best linear estimator, in which only the samples from the source are used and their correlations are the same. Again, the constant scaling factor from auto-correlation is removed to simplify the estimation process. So the moving average simply takes the most recent p samples and uses the average to estimate the current sample. Finally, by using only the samples from the source as the input to the best linear estimator, the need for computing cross-correlation and obtaining the samples from other sensors are eliminated.

VI. COMPARISION TO OTHER APPROACHES

As one can see from the previous section, the data estimation needed to recover the lost data packet incurs little power penalty. But the most exciting feature of our approach is that it is reactive. Conventional approaches like Forward Error Correction (FEC) or Automatic Repeat Request (ARQ) all require acknowledgement. Acknowledgement is needed in ARQ to let the sender know whether it needs to retransmit the packet. FEC eliminates the need for retransmission, but acknowledgement is still needed to let the sender know if the packet has been received at all by the destination, no matter it is correct or not. Acknowledgement is very expensive in terms of power. A typical acknowledgement packet used in IEEE802.11 system has at least 314 bits. The worst part is that an acknowledgement is needed for every data packet no matter the packet is received correctly or not. Even in wireless environment, where BER tends to be high, most of the time, a packet is received correctly. So a more reasonable approach would be doing nothing when there is no error. Action is triggered only when error happens. Our approach does exactly that. Most of the time, the packets are sent from the source to the destination without any error. There is no acknowledgment or data estimation. Therefore no power is wasted. Only when the destination does not get any data packet for some time, or simple end-to-end error detection shows the packet received is corrupted, data estimation will be triggered.

A FEC approach or a hybrid FEC/ARQ approach introduce carefully designed redundancy into a data link to improve the reliability of the link. This is of course at the cost of more bandwidth and more complexity. The latter leads to more power consumption. Our approach simply exploits the redundancy already in the network. So the penalty paid is much less. As shown from the simulation, simply using the previous sample can give fairly good estimation, the cost of which is almost zero.

Approaches, which choose to reduce or eliminate the redundancy inside a sensor network so that fewer bits can be transmitted from the source, can only save one or two bits per data packet. The complexity involved in implementing them is not justified, especially when compared to the saving of at least 304 bits per data packet brought by our approach with almost no cost.

VII. SIMULATION SETUP

Two sensor nodes were placed by about ten feet apart near a window at Berkeley Wireless Research Center. The two nodes collected data independently but at about the same time with same sampling frequency. Same numbers of samples were collected and were stored in a log file after the logging was finished. Two types of data were collected: temperature and light intensity. The light intensity data was collected around 2:30pm for an hour when the position of the sun is changing most at that time of the year. 3000 samples were collected at one-second interval. 20,000 samples of temperature data were also collected at same frequency for about six hours. In actual office building environment, temperature data is usually taken at a frequency of minutes and light intensity data is taken every 10's of seconds.

Four different estimators are compared, namely the best linear estimator using both sensor's data, last sample of the same sensor, moving average of the same sensor's data using the same number of samples and the best linear estimator only using the same sensor's data.

The simulation program is written in MATLAB. The program reads sensor data from the two data files logged at the same time by the two sensors. It implements a sliding window to compute the auto-correlation and cross-correlation functions. The simulation results are shown in Figure 1 to 6.

VIII. SIMULATION RESULTS

Figure 1 plots the light intensity measurements versus time. Four different estimates are also shown on the same figure. The estimator order p is 10 and the sliding window size W is 100. Figure 2 is similar to Figure 1 except that the temperature data is plotted instead. The temperature data is relatively flat, while the light intensity data varies a lot.

Figure 3 uses the same set of data as in Figure 1, but those data are down sampled at every 10 seconds so the correlation pattern will be similar to that of the office building environment. A close up snapshot is given for an interval of about 10 samples. It is clearly seen from the figure that the moving average estimate is the worst, while the correlation estimates are better than the last sample estimate except when the data is constant. This more easily seen in figure 4, where the estimation errors of different estimators are shown for the same interval.

Figure 5 and 6 plot the mean square errors divided by the current sample value for different p and W. W is 250 for all plots in Figure 5 and p is 2 in Figure 6. The samples used are the same as those in Figure 3. It can be seen in these two figures that the estimator using both sensors' data performs worse than the last sample estimator except for small p and large W. This is mainly due to the fact that the matrix to be inverted is ill conditioned for some samples. The estimator using the same sensor's data doesn't have this problem since Levinson-Durbin algorithm is used there, which is an order-recursive method. Similar thing can be done to solve the equations involving cross-correlations. The figures also show the simple estimator using only the last sample gives fairly good result.

VIIII. CONCLUSIONS

An ultra-low power MAC has been proposed in this paper. It is a distributed method. It is very simple and robust. It does not require any global or local synchronization. The main radio can be turned off whenever it is not receiving or transmitting. This low power design can be applied to any broadband wireless network where power is the key. For sensor networks, the inherent redundancy can be further exploited to greatly improve power efficiency. And it is shown that, by using very simple data estimation, a saving of at least 304 bits can be obtained per packet. This is compared to very complicated signal processing algorithms that remove redundancy directly. Those algorithms save only one to two bits per packet.

ACKNOWLEDGMENT

We would like to thank Dietrich Ho for setting up the sensor nodes to collect the data used in this paper and Prof. Michael Honig of Northwestern University for his valuable advice. We also acknowledge the contributions of the Pico-radio group at BWRC.

REFERENCES

[1] S. G. Fischer, T. A. Wysocki and H. J. Zepernick, "MAC protocol for a CDMA based wireless ATM LAN", *IEEE International Conference on Communications*, Montreal, Que., Canada, 8-12 June 1997.

[2] Y. Nakagawa, H. Uchiyama, H. Kokaji, S. Takahashi, M. Suzuki and M. Kanaya, "Multi-channel ad-hoc wireless local area network", *48th IEEE Vehicular Technology Conference*, Ottawa, Ont., Canada, 18-21 May 1998.

[3] K. M. Sivalingam, M. B. Srivastava and P. Agrawal, "Low power link and access protocols for wireless multimedia networks", *47th IEEE Vehicular Technology Conference*, Phoenix, AZ, USA, 4-7 May 1997.

[4] K. M. Sivalingam, M. B. Srivastava, P. Agrawal and J. Chen, "Low power access protocols based on scheduling for wireless and mobile ATM networks", *1997 IEEE 6th International Conference on Universal Person Communications Record*, San Diego, CA, USA, 12-16 Oct. 1997.

[5] S. Lee and D. Cho, "Distributed reservation CDMA for wireless LAN", *IEEE GLOBECOM '95*, Singapore, 13-17 Nov. 1995.

[6] D. Lopez-Rodriguez and R. Perez-Jimenez, "Distributed method for channel assignment in CDMA based ad hoc wireless local area networks", *1999 IEEE MTT-S International Topical Symposium on Technologies for Wireless Applications*, Vancouver, BC, Canada, 21-24 Feb. 1999.

[7] M. Joa-Ng and I. Lu, "Spread spectrum medium access protocol with collision avoidance in mobile ad-hoc wireless network", *IEEE INFOCOMM 1999*, New York, NY, USA, 21-25 March 1999.

[8] K. Sohrabi and G. Pottie, "Performance of a novel self-organization protocol for wireless ad-hoc sensor networks", *50th IEEE Vehicle Technology Conference*, The Netherlands, September 1999.

[9] J. Rabaey, J. Ammer, J. L. Silva Jr., D. Patel and S. Roundy "Pico radio supports ad hoc ultra-low power wireless networking", *Computer Magazine*, July 2000.

[10] J. Rabaey, J. Ammer, J. L. Silva Jr. and D. Patel, "PicoRadio: ad-hoc wireless networking of ubiquitous low-energy sensor/monitor nodes", *IEEE Computer Society Workshop on VLSI*, Orlando, FL, USA, 27-28 April 2000.

[11] Wireless LAN Media Access Control (MAC) and Physical Layer (PHY) Specification, *ANSI/IEEE Std 802.11*, 1999 Edition.

[12] S. S. Pradhan and K. Ramchandran, "Distributed source coding using syndromes (DISCUS): Design and construction", *Proc. IEEE Data Compression Conference (DCC)*, Snowbird, UT, USA, 29-31 March 1999.

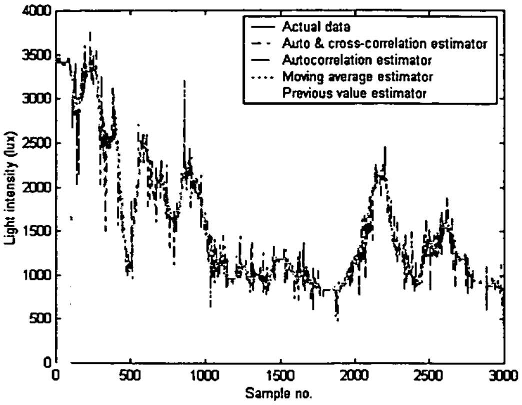

Fig. 1. Light intensity vs. time (sampled every second, W=100 & P=10)

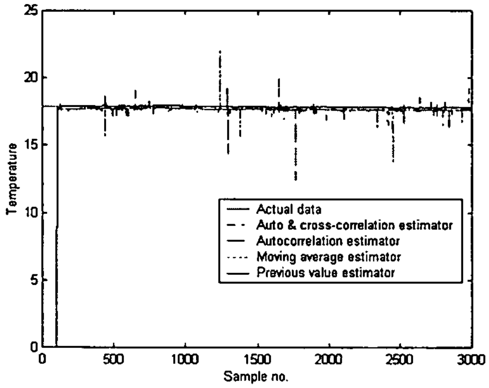

Fig. 2.Temperature vs. time (sampled every second, W=100 & P=10)

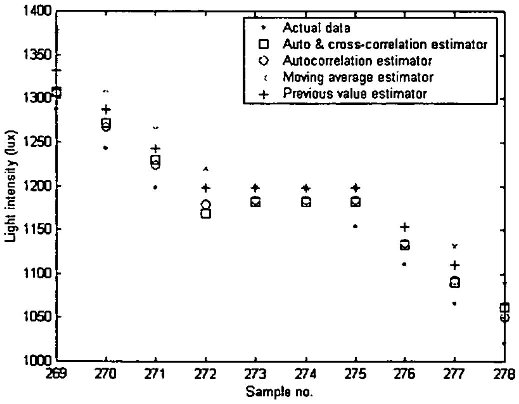

Fig. 3. Comparison of four different estimators (light intensity data sampled at every 10 seconds, W=250 & P=2)

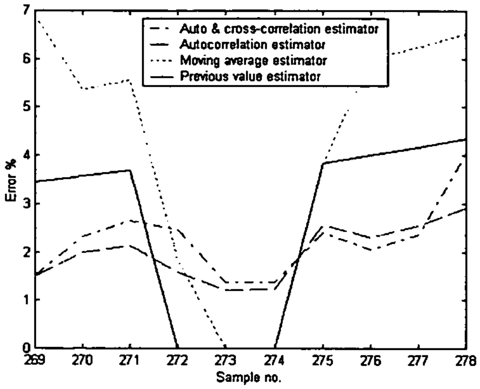

Fig. 3. Instantaneous error rate for light intensity data (sampled at every 10 seconds, W=250 & P=2)

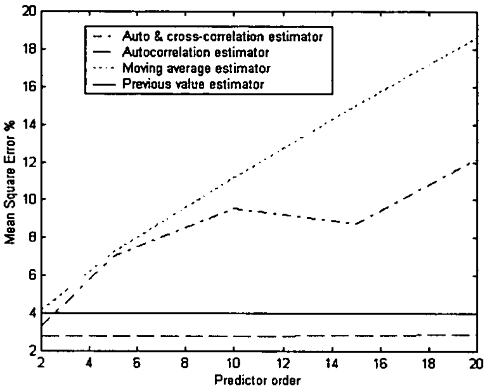

Fig. 4. Mean square error for different estimator orders (sampled at every 10 seconds, W=250)

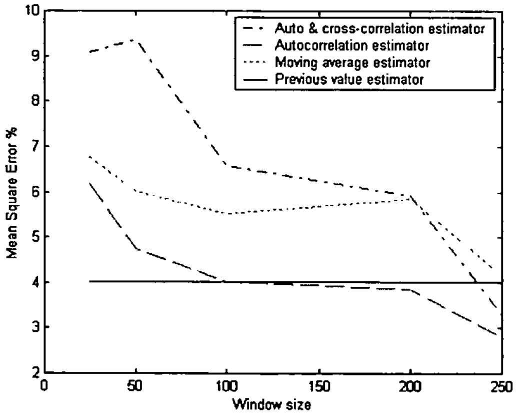

Fig. 6. Mean square error for different window sizes (sampled at every 10 seconds, P=2)